**Title**
Implementation of an iterative headway-based bus holding strategy with real-time information



**Journal**
Public Transport, 4(3)

**ISSN**
1613-7159

**Authors**
Chen, Qin
Adida, Elodie
Lin, Jane

**Publication Date**
2013-03-01

**DOI**
10.1007/s12469-012-0057-1

Peer reviewed

# Implementation of an Iterative Headway-based Bus Holding Strategy with Real-time Information

Qin Chen[1]. Elodie Adida[2] . Jane Lin[3]

**Abstract** In high frequency bus service lines, buses often come irregularly at the stops, often in bunches, due to the uncertainty of the passenger demand and behaviors, and the unexpected conditions on the roads. Vehicle holding is a commonly used strategy among a variety of control strategies in transit operation in order to reduce bus bunching and regulate bus headways. This paper investigates a control strategy of holding a group of buses at a single or multiple control point(s). By incorporating any possible passenger boarding activities during holding, a single control point problem is developed and extended to multiple control points to reduce the variance of headways for the downstream stops. The problem is a non-convex optimization programming with linear constraints that minimizes the total passenger waiting time both on-board and at stops. A heuristic is then developed that is easy and fast to implement, which makes it suitable for real-time implementation. The model is evaluated with a simulation case study by using the real-time bus operation data (i.e., Automatic Vehicle Location and Automatic Passenger Count data) from the Chicago Transit Authority (CTA). The simulation results show that considering the boarding activities in the total waiting time, our model mitigates the error propagation and maintains steady performance, compared to the common models in the literature, which do not consider boarding while holding.

***Keywords***: *Transit operations, Headway regularity, Bus holding, Control strategy.*

[1] Department of Civil and Materials Engineering. University of Illinois at Chicago. 842 W. Taylor Street, Chicago, IL 60607. Email addresses: qchen23@uic.edu

[2] University of California at Riverside, Anderson Graduate School of Management, 225 Anderson Hall, 900 University Ave., Riverside CA 92521. Email address: elodie.goodman@ucr.edu

[3] Corresponding author. Department of Mechanical and Industrial Engineering, and Institute for Environmental Science and Policy, University of Illinois at Chicago. 842 W. Taylor Street, Chicago, IL 60607. Email addresses: janelin@uic.edu

## 1 Introduction

Headway regularity is one of the commonly used transit vehicle on-time performance measures. If the vehicle is early or late, it likely results in very short headway with the preceding or the following vehicle, i.e., bunching (and at the same time a prolonged headway with the following or preceding vehicle). In either case, neither the transit agency nor the passengers are happy. Daganzo (2009) showed that on a high frequency bus route, without intervention, bus bunching was almost unavoidable regardless of the driver's or the passengers' behavior. In the Chicago Transit Authority (CTA) performance report, bus bunching is defined as a bus interval (time between two consecutive buses at a bus stop) that is 60 seconds or less. According to its monthly performance report in 2011 the CTA bus ridership was 44.3 million per month and there was an average of 2.3% bus bunching (of all weekday bus headways) system wide on weekdays. Most of the bunching occurred on high frequency and high demand routes, which typically serve the Central Business District area during peak hours, making the passengers' experiences with the transit service only less pleasant[4]. Bunching is one of the high priority issues transit agencies deal with on a daily basis.

Bus bunching occurs due to the inevitable varying running time between stops and the uncertainties associated with passenger arrival activities. Once a bus is behind the schedule, the gap with its leading bus becomes large, leaving more passengers to be picked up at downstream stops, which further delays the bus. Meanwhile, the successive bus collects fewer passengers at stops, which results in less dwelling time, and thus catches up, creating bus bunching; see, e.g., Bellei and Gkoumas (2010).

Generally speaking, there are two ways to solve buses bunching. One is to speed up the leading bus which is behind of schedule with stop skipping or short-turning by bypassing demand at some downstream stops. These strategies reduce the overall waiting time for some passengers but prolong the waiting time for skipped passengers. The other way is to slow down the following bus which is ahead of schedule, either by adjusting the cruise speed between stops or holding that bus at some stops (referred to as control points). However, cruising speed control is difficult to implement in practice due to the variable bus running time in continuously changing traffic conditions or extreme weather. Holding control refers to holding the vehicle at a control point for extra time after regular boarding and alighting activities. This strategy slightly increases the waiting time for passengers in the held vehicle, but the overall waiting time of all passengers on the route can be greatly reduced. Shen and Wilson (2001) found that holding strategies could save 10%-18% of passenger waiting time. The easy implementation of such a control method also makes it an attractive policy to transit agencies to improve on-time performance.

In this study, we present a fast implementable real-time holding strategy for multiple buses at one or multiple control point(s). One of the important features of our proposed holding

[4] It is interesting to note that in 2009, the CTA had a slight drop in bus bunching down to 2.5% due to budget cut which resulted in less frequent service.

policy is its fast computation which makes it suitable for real-time implementation in a bus system. It is based on an online algorithm that continuously updates the holding policy with continuous streams of real-time Automatic Vehicle Location (AVL) and offline Automatic Passenger Count (APC) data. With availability of AVL and APC data, real-time bus operation data has become available continuously for transit service monitoring and assessment. This presents an opportunity to implement many operational policies including holding policies in real time and online. As we will demonstrate later in the paper, this holding algorithm represents considerable improvements over the common ones available in the current literature.

The paper is organized as follows: Section 2 gives a brief literature review, details the assumptions and describes the model. Section 3 includes a description of the real-time CTA operation data set, followed by a simulation case study demonstrating the effectiveness and limitations of the approach in Section 4. Lastly, Section 5 summarizes the findings and draws the conclusions.

## 2 Literature Review on Transit Vehicle Holding Studies

Holding strategies can be classified into schedule-based holding and headway-based holding. In the schedule-based holding strategy, if a vehicle arrives early at the control point, it is held until a target scheduled departure time is reached. If the vehicle is already late, no holding takes place (Oort and Wilson, 2010). The schedule-based holding relies on the pre-specified static schedule, which is not desirable when buses are running on short headways, i.e., schedule becomes irrelevant from the passengers' point of view. In such cases, a headway-based holding strategy is more appropriate.

The headway-based holding aims at reducing the headway variability to equalize the headways among the controlled vehicles because the average waiting time decreases as the variance of headway reduces (Welding 1957). Headway-based holding is usually applied in frequent service routes with a short dispatching headway (Hickman, 2001; Fu and Yang, 2002; Bellei and Gkoumas, 2009). High frequency service takes place in busy city centers and during peak hours. Thus the headway-based holding strategies may be of higher impact and are the focus of this paper especially relevant to the bus service in the Chicago metropolitan region.

Headway-based holding problems are formulated differently in the literature. One approach is based on the average waiting time as a function of the mean and variance of headway, and simulating the distribution of headway and duration of disruption using historical data (Barnett, 1974; Turnquist, 1981; Abkowitz and Tozzi, 1986; Abkowitz and Lepofsky, 1990). This approach addresses the stochastic effects of the elements such as running time and passenger arrival rate. It only accounts for waiting time for downstream at-stop passengers and ignores the on-board passenger waiting-time, which may not be negligible.

Alternatively, optimization models with the objective to minimize the passenger waiting time (either only at-stop waiting time or a combination with on-board waiting time) are popular. This approach can model the passenger waiting time sufficiently accurate and handle multiple holdings. However, currently this type of model is difficult to solve for complex problems (Xuan et al., 2011). Some works use existing heuristics that may not render near-optimal solutions with less restrictive assumptions fast enough for real-time application (Cortes et al., 2010), and others make stronger assumptions to simplify the models so that readily available commercial packages such as MINOS can apply (Delgado et al., 2009). In this study, our objective is to develop a headway-based holding optimization model that is fast, easy and cheap to implement, and suitable for real-time applications.

Boarding while holding is an activity that has been ignored in optimization modeling (e.g., Puong and Wilson, 2008; Sun and Hickman, 2008), because otherwise the problem would become non-convex and hard to solve. While that simplification has greatly reduced the complexity of the bus holding problem, the literature often possesses an inconsistency that pertains to how to treat holding time in relation to dwell time, headway and subsequently the number of passengers boarding and alighting at the control point. Dwell time at the control point is defined as the sum of the regular dwell time determined by the passenger arrival rate and service rate and the bus holding time during which no passenger boarding and alighting are considered. The number of on-board passengers, on the other hand, is calculated inclusive of possible boarding and alighting activities during holding. Such inconsistency may become problematic especially when the demand is high at the stop and the holding time is relatively long. Furthermore, as we will demonstrate later in this paper, the effect of boarding while holding on waiting time saving cannot be ignored. Recently there have been a few studies explicitly taking boarding while holding into account with simulation approach. For example, Bellei and Gkoumas (2009) evaluated, by using Monte Carlo simulation, a combined strategy of holding and transit signal priority to reduce headway variation. Toledo et al. (2010) used simulation software to evaluation the stochastic effects of bus holding strategy, with explicit handling of boarding while holding in the microsimulation.

In this study, a workable optimization method with high efficiency heuristics is presented for real-time headway-based holding control. A modified single control point bus holding problem is formulated by considering the possible passenger boarding activities during holding. This significantly increases the complexity of the problem, due to the induced interdependence of variables among vehicles and stops. Therefore, the problem is non-convex and thus difficult to solve exactly. A heuristic is thus presented to solve the problem considered, We also demonstrate how the proposed holding strategy, designed for a single control point holding, can be extended to multiple control points with stochastic running time taken into consideration. Lastly, the model is evaluated in a case study by using the real-time bus operation data (i.e., automatic vehicle location (AVL) and offline automatic passenger Count (APC) data) from the CTA. While real-time holding problems have been studied for city rail operations (O'Dell &Wilson, 1999; Ding and Chien, 2001; Eberlein et al., 2001), few studies have evaluated real-time bus holding strategies. This is one of the few studies that emphasize on real-time implementability of a bus holding strategy. In summary, this paper differs from the preceding approaches in the following ways:

- We explicitly consider the passenger boarding activities during holding, reducing the waiting time for those who arrived during the holding in an optimization model;
- Although the model makes the assumption that running time is deterministic, the

implementation releases this assumption and considers the running time as stochastic;

- Real-time bus operation data is obtained to test both single control point and multiple control points holding strategy.

## 3 Problem Formulations

In this section, we propose a more realistic bus holding strategy than those found in the literature where passenger boarding activity during holding is typically ignored while in reality passengers do continue boarding during holding. We first formulate the problem (denoted as ***P***) of holding a group of buses at a predefined control point (denoted as stop ***k***) and later extend the strategy to multiple control points.

To begin, we start with the key assumptions necessary for the proposed formulation, followed by the formulation itself. We then present a heuristic algorithm to approximate the solution as the problem, as it turns out not surprisingly, is a non-linear, non convex one. The convergence issue of the heuristic approach will be discussed. Finally in this section, we show how to implement this algorithm for multiple control points.

### 3.1 Assumptions

We make the following assumptions:

1) The bus running time between stops and the service rate at each stop are assumed to be deterministic in the model formulation. Note, however, that this assumption can be relaxed in the implementation as we will show later in the case study.
2) Passenger boarding during holding follows the same arrival rate as that during the regular dwell time. Furthermore, passenger arrivals are assumed to be uniform. This assumption may not be very far from the reality given that in a headway-based (high frequency) service passengers tend to arrive approximately evenly as compared to a timetable based (low frequency) service where passengers tend to time their arrival depending on the schedule.
3) Dwell time is an affine function of the number of boarding passengers, i.e., boarding is assumed to be the dominant activity, in a front-on rear-off boarding/alighting setting.
4) The number of alighting passengers from a bus at a given stop is proportional to the number of passengers on board at the time of arrival.
5) No bus overtaking is allowed. This is a common simplification seen in the literature (e.g., Sun and Hickman, 2008).
6) No vehicle capacity constraint is considered. Considering capacity constraints is one of the causes to bus bunching. However, in reality especially in the U.S. cities bus capacity is hardly reached unless there is a sudden surge in ridership or service disruption. This is evident by the bus passenger count data provided by the CTA. In the case of sudden surge in ridership or service disruption, the bus holding strategy itself would not help to regulate the service and other strategies such as stop-skipping would have to be considered as well, which is beyond the scope of this paper.

### 3.2 Problem *P*

In problem (***P***), the objective is to determine the holding time for each bus ***i*** in the bus set (denoted ***Im***) at stop ***k*** so as to minimize the total passenger waiting time both on-board and at the downstream stops. The impacted downstream stops on the considered route from the impacted stop correspond to set ***In={k,k+1,k+2,…}*** and an impacted stop is denoted ***j∈In***. The rest of the notations are listed below:

$r_j$ ---deterministic passenger arrival rate at stop ***j***;

$q_j$ ---alighting fraction of the total number of on-board passengers at stop ***j***;

$A_{i,j}$ ---number of alighting passengers off bus ***i*** at stop ***j***;

$B_{i,j}$ ---number of boarding passengers onto bus ***i*** at time of departure from stop ***j***;

$b_{i,j}$ ---number of passengers boarding bus ***i*** at the end of the regular dwell time at stop ***j***; observe that $b \leq B$ at the control point, and *b=B* at other stops;

$L_{i,j}$ ---passenger load on bus ***i*** at time of departure from stop ***j***;

$R_j$ ---deterministic running time between stops ***j-1*** and ***j***;

$\alpha, \beta$ ---parameters in determining dwell time as a function of boarding;

$S_{i,j}$ ---regular dwell time for bus ***i*** at stop ***j*** (i.e., no holding time);

$t_{i,k}$ ---holding time for bus ***i*** at control point ***k***;

$a_{i,j}$ ---arrival time of bus ***i*** at stop ***j***;

$d_{i,j}$ ---departure time of bus ***i*** at stop ***j***: $d_{i,j} = a_{i,j} + S_{i,j}$ if $j \neq k$ and $d_{i,k} = a_{i,k} + S_{i,k} + t_{i,k}$.

The problem (***P***) of multi-bus holding at a given control point (stop) ***k*** can be formulated as follows:

$$\min Z = \sum_{In}\sum_{Im}(a_{i,j} - d_{i-1,j})^2 \times r_j / 2 + \sum_{Im} t_{i,k} \times (L_{i,k} - r_k \times t_{i,k}) + \sum_{Im} t_{i,k}{}^2 \times r_k / 2 \qquad (1)$$

$$\text{St.} \quad S_{i,j} = \alpha + \beta \times b_{i,j}, \forall i \in Im, j \in In \qquad (2)$$

$$A_{i,j} = q_j \times L_{i,j-1}, \forall i \in Im, j \in In \qquad (3)$$

$$B_{i,j} = r_j \times (d_{i,j} - d_{i-1,j}), \forall i \in Im, j \in In \quad (4)$$

$$b_{i,j} = B_{i,j}, \forall i \in Im, j \in In, j \neq k \quad (5)$$

$$b_{i,k} = r_k \times (d_{i,k} - t_{i,k} - d_{i-1,k}), \forall i \in Im \quad (6)$$

$$d_{i,j} = a_{i,j} + S_{i,j}, \forall i \in Im, j \in In, j \neq k \quad (7)$$

$$d_{i,k} = a_{i,k} + S_{i,k} + t_{i,k}, \forall i \in Im \quad (8)$$

$$a_{i,j} = d_{i,j-1} + R_j, \forall i \in Im, j \in In \quad (9)$$

$$d_{i,j} \leq a_{i+1,j}, \forall i \in Im, j \in In \quad (10)$$

$$L_{i,j} = L_{i,j-1} + B_{i,j} - A_{i,j}, \forall i \in Im, j \in In \quad (11)$$

$$t_{i,k} \geq 0, \forall i \in Im \quad (12)$$

$$d_{i,j} \geq 0, \forall i \in Im, j \in In \quad (13)$$

$$a_{i,j} \geq 0, \forall i \in Im, j \in In \quad (14)$$

$$L_{i,j} \geq 0, \forall i \in Im, j \in In \quad (15)$$

In the objective function, both at-stop and on-board passenger waiting times due to holding are considered. The first term in expression (1) is the summation of the total at-stop passenger waiting time since the departure of the preceding bus (***i-1***) till the arrival of the current bus (***i***), for all buses in the impact set ***Im*** and at all impacted stops; the second term is the total regular on-board waiting time (from arrival to departure) without considering the additional passengers who boarded during holding; and the last term is the total additional on-board waiting time for those additional boarding passengers during holding.

Constraints (2) and (3) define dwell time and passenger alighting, respectively. Constraints (4) to (6) define boarding separately for the control stop (i.e., stop ***k***) and the non-control stops (i.e., the downstream impacted stops ***j∈In*** and ***j≠k***). Constraints (7) and (8) define departure times at the non-control stops and the control stop, respectively. Given the deterministic running times assumption, constraint (9) defines the arrival time at each of the impacted stops and (10) ensures that the arrival time of bus ***i+1*** at stop ***j*** can never precede the departure time of bus ***i*** at stop ***j***, i.e., no overtaking is allowed. Constraint (11) states that the load when the bus departs from stop ***j*** equals to the load when it departs from stop ***j-1*** modified by the boarding and alighting activity at stop j. Lastly, (12)-(15) are non-negativity constraints on the variables.

Writing the objective function (1) into the matrix form, we have

$$\min Z = X^T MX + q^T X + C \tag{16}$$

where $\boldsymbol{X}$ is a vector defined by auxiliary variables $a_{i,j}$, $d_{i,j}$, and $L_{i,j}$ for all the buses and stops in the impact set, as well as the decision variable $t_{i,k}$; $\boldsymbol{M}$ is a matrix of coefficients that involve the quadratic terms and $\boldsymbol{q}$ is a vector of coefficients corresponding to the linear terms while $\boldsymbol{C}$ is a constant. Note that matrix $\boldsymbol{M}$ is not a positive semi-definite matrix.

The problem formulated here is thus a non-linear program with a non-convex quadratic objective function subject to linear constraints. The inclusion in the formulation of boarding activities while holding does not make the quadratic term matrix $\boldsymbol{M}$ a positive semi-definite matrix. Clearly, there is no closed form solution or simple way to compute an exact solution. To solve it, we have developed a heuristic algorithm which we describe below.

### 3.3 The Heuristic *P0*

The idea is to transform the original non convex problem (***P***) into a convex one and solve it iteratively and heuristically. To achieve this, an iterative holding time variable $\boldsymbol{T_{i,k}(p)}$ is introduced as the current holding time (of bus ***i*** at stop ***k***) in iteration p to approximate the exact holding time solution $\boldsymbol{t_{i,k}}$. Specifically, here are the steps in the proposed heuristic:

**Step 0:** Initialization: iteration number $\boldsymbol{p=0}$, and initial holding time $\boldsymbol{T_{i,k}(0)=0}$;

**Step 1:** Iteration: at iteration $\boldsymbol{p}$, combine the last two terms in the objective function, and rewrite the objective function as

$$\min Z = \sum_{In}\sum_{Im}(a_{i,j} - d_{i,j})^2 \times r_j / 2 + \sum_{Im} T_{i,k}(p) \times (L_{i,j} - r_k \times t_{i,k} / 2) \tag{17}$$

Notice that by changing the decision variable $\boldsymbol{t_{i,k}}$ to an iterative term $\boldsymbol{T_{i,k}(p)}$, the revised objective function as shown in expression (17) becomes convex and the optimization problem is easy to solve.

**Step 2:** $\boldsymbol{p=p+1}$, $\boldsymbol{T_{i,k}(p+1)= t_{i,k}}$, and repeat Step 1 until the convergence of decision variable $T_{i,k}(p)$ is achieved when $\left|t_{i,k} - T_{i,k}(p)\right| \le \varepsilon, \forall i \in Im$ (where ε is a predefined tolerance level).

This heuristic ***P0*** gives a solution provided that the algorithm converges. In addition, the resulting solution from ***P0*** need not be an optimal solution to problem (***P***). In the next subsections, we investigate the convergence of ***P0*** in Section 3.4 and optimality of its solution in Section 3.5.

### 3.4 Convergence discussion for *P0*

For the purpose of simple presentation and illustration, we write the optimization problem involved in ***P0*** in the following matrix form:

$$\min Z = X^T M_h X + q_h{}^T X + C \tag{18}$$

$$AX = b \tag{19}$$

$$X \geq 0 \tag{20}$$

Comparing (18) with (16), it is observed that $\boldsymbol{M_h}$ is a matrix that involves the arrival time and departure time coefficients, while $\boldsymbol{M}$ is a coefficient matrix which corresponds to all the components in the decision variable; $\boldsymbol{q_h}$ is only associated with holding time and passenger load while $\boldsymbol{q}$ contains all the components in the decision variable.

We then derive the Karush–Kuhn–Tucker conditions:

$$2M_h X^{(p+1)} + q_h{}^{(p)} + A^T V^{(p+1)} + U^{(p+1)} = 0 \tag{21}$$

$$U^{T(p)} X^{(p)} = 0 \tag{22}$$

$$X^{(p)} \geq 0 \tag{23}$$

Notice that in a given iteration *g*, $\boldsymbol{q_h}$ can be written as $q_h{}^{(g)} = q_h^0 + Q_h X^{(g)}$, where $\boldsymbol{q_h}$ is a constant vector. Then equation (21) becomes

$$2M_h X^{(p+1)} + Q_h X^{(p)} + A^T V^{(p+1)} + U^{(p+1)} + q_h = 0 \tag{24}$$

As it turns out, the connection between iterations in equation (24) is not obvious because matrix $\boldsymbol{M_h}$ in iteration *p+1* only involves the non-zero coefficients of departure time and arrival time, while the matrix $\boldsymbol{Q_h}$ in iteration *p* only involves the non-zero coefficients of holding time and passenger load. Furthermore, $\boldsymbol{M_h}$ is not strictly diagonal and it is not desirable to break the matrix into sub-matrices because it will create more off-diagonal non-zero terms and add additional complexity to the formulation. Thus, proving that the heuristic converges is not straightforward. However, in numerical experiments, the heuristics showed fast convergence to a solution.

To illustrate the convergence of the heuristic, we set up a large number of simulation runs for various case studies. Up to 100 scenarios with a variety of input parameters and problem size (number of controlled buses ***m*** and number of impacted stops ***n***) have been tested in the simulations. Randomly generated running times between stops based on the statistics in the observation data[5] and fixed passenger arrival rates (1.7 persons per minute, also obtained from the CTA data set which will be described later) were used in the simulation. For each scenario,

[5] The parameters were derived from real time bus route operation data provided by the CTA.

we repeated the simulation run with randomly generated starting points. With the convergence error $\varepsilon$ set to be 1.0e-5, we recorded the CPU time as well as the number of iterations.

The numerical results suggest that the heuristic does converge to a solution. Table 1 presents selected results from eight randomly chosen sample runs among all the scenarios tested. As shown in Table 1, ***P0*** converges fast and requires a small number of iterations. For example, in the first sample run, the absolute difference between the starting point and the next solution is 19.4; after 5 iterations, the absolute difference drops to 9.69E-07, less than the given tolerance $\varepsilon$ =1.0e-05. An important feature of the proposed algorithm is its fast computing time at each given control point, even when the problem scale is comparatively large, i.e., with large number of controlled buses and impacted stops. For example, as the fourth sample run shows in Table 1, the algorithm converges in 4 iterations in 0.7108 seconds with a problem size of ten controlled buses and ten impacted stops. Moreover, by comparing the two sample run results for the same problem size (e.g. sample run 1 and 2), we notice that by carefully selecting the starting point for the search, the algorithm converges even faster with less number of iterations.

[Table 1]

Although it is difficult to mathematically prove theoretically the convergence of our heuristic algorithm, numerical implementation indicate the good practical performance of the heuristic method for a large range of instances of the problem considered. While some off-the-shelf commercial software may be able to solve small instances of the non convex problem, the performance of this software sharply decreases as the problem size increases, while our heuristic method is based on solving convex quadratic programs and thus continues to perform well even for large size of the problem. The fast computing time make it possible for online real-time implementation of bus holding strategies, especially when the problem size is large.

### 3.5 Implementation of *P0* and multiple control points problem

The problem (***P***) assumes deterministic bus running times; however, in reality bus running time is stochastic with large variance due to the road conditions and drivers behavior. Normal and lognormal distributions are commonly used to represent the stochastic effects of travel time. According to Herman and Lam (1974) and Turner and Wardrop(1951), shorter trips tended to be normally distributed while longer trips followed log-normal distribution. In our study, the trips between stops are relatively short in the urban area. Therefore, bus running time $R_j$ is assumed to follow normal distribution with mean $mR_j$ and variance $varR_j$. It is worth noting that lognormal distribution can be applied to the longer trips in the similar fashion without losing the generality of our algorithm's properties. Then we implement the algorithm in a rolling-horizon concept similar to Eberlein et al. (2001) as follows:

**Step 1:** predefine the number of impacted buses $m$, the number of impacted stops $n$, and the control point $k$ and the first controlled bus $I$;

**Step 2:** When bus ***I*** arrives at the control point ***k***, the arrival time of the impacted buses at the control point, the departure time of the preceding bus at the impacted stops, and passenger loads of the impacted buses before arriving at the control point are captured as input, either by updating real-time information or prediction based on existing knowledge. Then the deterministic running time $mR_j$ is used to run the algorithm ***P0*** and get the optimal holding time for the impacted bus group {***I, I+1,…,I+m-1***}; only the holding time for bus ***I*** at the control point ***k*** is implemented at this point;

**Step 3:** With optimal holding time calculated in step 2, we use stochastic running time to predict the unknown arrival and departure time for the downstream stops and the further bus trips. Project the trajectories and calculate the downstream headways for evaluation purposes;

**Step 4:** When the next bus ***I+1*** arrives at the control point ***k***, repeat steps 2 and 3 with the updated input values including the new arrival rate as well as the prediction results from step3.

It is note-worthy that the above single control point algorithm can be extended to multiple control points by pre-defining multiple control points and implementing the algorithm to these points. The motivation for holding at multiple control points is as discussed in Abkowitz and Engelstein (1984) and Sun and Hickman (2008): the holding effect typically dissipates quickly at the downstream stops, making the single point holding inefficient.

## 4 Case-study of real-time bus holding implementation

This section presents a simulation case study to demonstrate the implementation of the proposed bus holding policy in a real time environment. The case study uses real world bus location and count data recorded on a bus route, which runs north-south across the Chicago downtown area and has a relatively high volume of passenger boarding and alighting during peak hours. In addition, we compare our model with a similar study in the literature (Fu and Yang, 2002) with the data used in that study. The results are presented in this section.

### 4.1 Real-time Bus data from Chicago Transit Authority

The CTA serves the city of Chicago and the surrounding suburbs with more than 150 routes, with downtown Chicago generating the largest demand for bus service. All CTA buses are equipped with the AVL and APC systems. In this study, we use the same data set as used in Lin and Ruan (2009). As can be seen in Figure 1, the study route is a CTA route on a major urban street in the west side of downtown Chicago. The north-southbound route intersects with many other bus routes and connects to the subway system, resulting in a relatively high ridership. It takes about 80 minutes to run on the 14 mile route and there are about 110 stops and 13 time points in total, which are planned geographic locations along the route for schedule adherence control purpose. In other words, these time points are the potential control points for bus holding. Geographically, time points are often physical stops in a bus route. Therefore, the time points were used instead of stops in our case study.

[Figure 1]

One month of original weekday AVL and APC records in September 2006 contained bus operational events (e.g., serviced a stop, dwell time, passenger count, arrival time, departure time) at the time-point level. Using the bus trip records in both directions during the morning peak hour (from 7:00AM to 10:00AM), we compute and calibrate the following model parameters: arrival rate $r_j$ (=1.04~ 3.06 persons per minute from stop to stop), alighting fraction $q_j$ (=0.06), mean and variance of bus running time ($mR_j$ and $varR_j$) between any two consecutive stops $j-1$ and $j$ (see Table 2), and the parameters in dwell time-boarding function $\acute{a}$ (=0.05) and $\hat{a}$ (=0.08) in Eq. (2) by fitting a linear regression model based on the dwell time and boarding counts. The average observed headway was 7 to 11 minutes, which falls in the range of headway-based control strategy in the literature. Lastly, as shown in Figure 2, the passenger load exceeding 50 during the peak hours accounted for 4% of all the bus trips during the peak hours. With the fact that most of the transit buses have a capacity of 50-70 in the US and Canada[6], it is not unreasonable to assume no vehicle capacity constraint in this study.

[Figure 2]
[Table 2]

The simulation was built on the southbound bus trip records during the morning peak hours on Wednesday, September 13, 2006 from 7:00AM to 10:00AM. There were 12 buses running in the route with 7-11 min headways during the peak hours.

### 4.2 Performance Measures

Following other studies in the literature (e.g., Fu and Yang, 2002; Sun and Hickman, 2008), this study uses the following metrics to evaluate the effectiveness of the holding strategy:

1) *Percent time saving* compared to the no holding scenario is defined as the change in the objective value (passenger-minutes) between the holding and no-holding scenarios divided by the objective value without holding. The more time saved, the more passengers benefit from the holding strategies.
2) *Holding time* for bus $i$ (minutes): generally bus should not be held too long in a bus holding strategy.
3) *Average holding time* at a control point (minutes per bus). When there are $m$ buses in the

[6] Transit Capacity and Quality of Service Manual, 2nd Ed.

impacted bus set, the average holding time at this control point is calculated by taking the mean value of the $\boldsymbol{m}$ individual holding times.

4) *Percent improved headway stability* is derived with the following equation:

$$(CV(headway)_j^{nohold} - CV(headway)_j^{hold}) / CV(headway)_j^{nohold}$$

where coefficient of variation *CV(headway)* is the ratio of the standard deviation and the mean value of headway, which is determined among the impacted buses in set $\boldsymbol{Im}$, for a downstream stop $\boldsymbol{j}$ (here we only consider $\boldsymbol{j \in In}$). This metric measures how stable (or constant) the headways are before and after holding. A larger value indicates better improvement in reducing the headway variance, which is the objective of holding.

5) *Total computing time* (CPU seconds) is based on the run time in Matlab on a 2.2GHz Intel Core 2 Duo CPU computer.

### 4.3 Simulation results

#### *4.3.1 Performance of the proposed model*

Applying the model (***P***) to the CTA bus operation data, we set the number of impacted buses $\boldsymbol{m}$=3 and stops $\boldsymbol{n}$=3, and the control point at stop 7 for its relatively high demand (the arrival rate at stop 7 is 3.06 passengers per minute); hence the impacted stops are {*8, 9, 10*}. We applied the control strategy from bus *2* to the following buses, and obtained three runs in this example to get the optimal holding times for bus {*3, 4, 5*}. The inputs to each run and the results are summarized in Table 3. The input parameters, i.e., impacted bus arrival times at the control point {$\boldsymbol{a_{i,7}}$}, previous bus departure times at the control point and the impacted stops {$\boldsymbol{d_{1,j}}$}, and the impacted bus loads before control points {$\boldsymbol{L_{i,6}}$}, are obtained either from existing data or prediction data based on current information. Finally, the arrival rates and the running time are read from the statistic profile as mentioned in Section 4.1.

[Table 3]

In the first run, with six iterations and CPU computing time of 0.347 seconds, the simulation gives a solution of holding time for buses {*2, 3, 4*} to be {*5.961, 6.873, 8.728*} minutes respectively at the control point of stop 7. The average time saving compared to no-holding case is 7.13%. Bus 2 is held at stop 7 for 5.961 minutes. The subsequent arrival times at stop 7 for buses *3, 4, 5* and the departure time from stop 7 for bus 2 are predicted by using the stochastic running time $\boldsymbol{mR_j}$ and $\boldsymbol{varR_j}$. The updated arrival and departure information then become the new inputs to the next model run and so on. The final implemented holding times for the impacted buses {*2, 3, 4*} are {*5.961, 3.309, 6.803*} minutes, respectively, and as a result the average time saving is 13.29%.

The bus trajectories for the cases with and without holding are presented in Figure 3(a) and (b), respectively. As shown in Figure 3(a), when there is no holding, buses 1, 2 and 3 bunched together at various stops along the route. With the proposed single bus holding strategy implemented at stop *7*, the headways between buses 1, 2 and 3 become more even at the downstream stops from stop 7, as seen in Figure 3(b). It is worth noting that in the single holding strategy, the headway between buses 4 and 5 becomes noticeably smaller than that of no-holding (Figure 3(a)). The reason is that Bus 4 has light on-board passengers compared to Bus 5, and by applying a holding time to Bus 4, part of the downstream demands for Bus 5 were transferred to Bus 4, preventing further delay of Bus 5.

[Figure 3]

Figure 3(c) shows a scenario in which holding was applied to bus set {*2, 3, 4*} at two control points, stops 3 and 7, both of which have high demand and are not located very close to each other. Comparing with the single holding strategy in Figure 3(b), multiple holding makes the headways between all the buses considered more even throughout the entire bus route, eliminating bus bunching between buses 1, 2, and 3. Additional differences in smoothing bus headways between multiple holding and single holding are summarized in Table 4.

[Table 4]

As shown in Table 4, in the multiple holding strategy, the average holding time at stop 7 is 5.01 minutes, slightly less than that in the single holding strategy. The total time saving is also similar, indicating comparable performance between the two in those two measures. However, multiple holding has much better headway stability in the subsequent stops, which is consistent with what is observed in Figure 3 (c). The above results suggest some benefits of multiple holding over a single holding strategy.

### *4.3.2 Advantages of considering boarding activities while holding*

To demonstrate any possible advantages of considering boarding activity in our model over the models in the literature that typically do not consider boarding activities while holding (referred to as "traditional model" hereafter), we re-constructed the traditional model (***T***) and formulate it in a structure similar to our proposed model by modifying the objective function in (1) to the following in (22). The difference exists in not including passenger activities while holding:

$$\min Z = \sum_{In}\sum_{Im}(a_{i,j} - d_{i\text{-}1,j})^2 \times r_j / 2 + \sum_{Im} t_{i,k} \times L_{i,k} \qquad (22)$$

Additionally, constraint (6) states the relationship between the number of boarding passengers and the associated headways at the control point. To be compatible with the traditional model, we have modified constraint (6) to the following:

$$b_{i,k} = r_k \times (d_{i,k} - d_{i-1,k}), \forall i \in Im \qquad (23)$$

Therefore, under the constraints (2)-(5) and (7)-(15), together with constraint (23), the new model is a convex quadratic problem, which can be readily solved using a standard optimization software package. To be consistent, we again use Matlab to code the model (***T***).

To compare the performance of the two models, we applied the multiple holding strategy to the same chosen CTA route as described in Section 4.1 within the two models separately. Again we considered 13 stops on the route. We constructed four scenarios of different numbers of impacted stops = *2, 3, 4, 5* respectively, while keeping the number of impacted buses at four. Table 5 summarizes the arrival rate at each stop and the lists of control points along the route in each scenario (Y indicates the location of the control points).

For demonstration purposes, the first control point is set at stop 2 or 3; the interval between the two successive control points is equal to the number of impacted stops. For example, as shown in Table 5, in the last scenario where the number of impacted stops(***n***)=5, the control points are set at stops {2, 8} with the associated arrival rates of 1.04 and 1.33 passengers/minute respectively; five downstream impacted stops following the control points are *{3,4,5,6,7}* and *{9,10,11,12,13}*.

[Table 5]

For both models (P) and (T), we computed the average holding time and the improved headway stability at the subsequent stops of the control point. We also computed the total waiting time, which is the objective function (1) for model (P) and (22) for model (T). Table 6 summarizes the average holding time for both models, and presents the following performance measures between models (P) and (T): the total waiting time saving in terms of passenger-minutes and the percentage of improved headway stability. Because waiting time is defined differently in the two models, they could not be compared directly. Thus we calculated the waiting time saving as such: in step one, we solved the traditional model(T) for a set of holding times {T1}and then fed {T1} to model (P) to obtain the waiting time WP({T1}); in step two, we solved model (P) for another set of holding time {T2} and the corresponding waiting time WP({T2}). The waiting time saving is then defined as WP({T1})-WP({T2}), where WP({T1}) represents the passenger waiting time that would have been like without considering boarding while holding, and WP({T2}) represents the passenger waiting time after boarding while holding was incorporated.

[Table 6]

As shown in Table 6, the average holding times in the two models are close to each other at many stops except a few with relatively high demands (e.g., in scenario n n=4, the holding time for models (P) and (T) at stop 7, which has the highest passenger demand, are {6.16, 7.21} minutes respectively, representing a whole minute (63 seconds) of difference in holding time. That is quite significant considering a typical dwell time at a CTA bus stop is usually less than 30 seconds.

In Table 6, the positive values in difference between models (P) and (T) indicate that (P) outperforms (T) by 1) saving more on passenger waiting time, and 2) delivering better improved headway stability (from the baseline, i.e., without holding) at the downstream stops. In other words, incorporating the extra boarding activities during holding in the total waiting time mitigates the error propagation and stabilizes the performance at downstream stops, and therefore saving more passenger waiting time in general. The total waiting time saving of (P) from (T) can be as large as 43 passenger minutes, and the difference in improvement of headway stability as high as 10% at high demand points (e.g., stop 7).

Given the fact that the control points are more likely to be set at the stops with relatively high demand (Fu and Yang,2002, Hickman, 2001) , our proposed model outperforms the traditional model especially on high demand scenarios. Also it is worth mention that in reality, it is better to over lay the control points with the impacted stops (i.e., making the control point interval less than the number of impacted stops) to increase the benefit of holding with improved headway regularity.

Moreover, we can see from Table 6 that the effect of holding becomes significant as the number of impacted stops increases. For example, in scenarios n={2,4,5} where the first control point is all set at stop 2, it is found that there is less holding time and more waiting time saving as the number of impacted stops (n) increases. The subsequent control points are not comparable as any control actions at the upstream stops could change the trajectory at the downstream stops. Similar findings are confirmed in the sensitivity analysis to be presented in Section 4.3.3.

#### *4.3.3 Sensitivity analysis for single control point problem*

The first sensitivity analysis concerns the bus holding performance with respect to the number of impacted stops and the number of impacted buses. The experiment was carried out as follows: we created a total of twenty-five scenarios (simulations) by varying the number of impacted buses and stops from *1* to *5,* respectively; the arrival rate at all stops was set at 1.7 persons per minute and the other input parameters were to the same as the ones in Section 4.3.1.

In Table 7, we notice that as the number of impacted stops increases, the average holding time decreases, as expected, from 5.4 minutes to 3.3 minutes because more at-stop waiting time is considered. Consequently, the percent of time saved increases as seen in Table 8, which suggests that as many impacted stops as possible should be included in the holding strategy analysis to maximize the benefits. On the other hand, the benefits are diminishing when the impacted stops are beyond four in this case study. Therefore, no more than four stops should be included in the impact set. Similar findings are observed in the number of impacted buses. The practical implication is that the impact sets (buses and stops) must be chosen carefully to maximize the effectiveness of the holding strategy.

As shown in Tables 7 and 8, the number of impacted buses does not affect the average holding time significantly. However, the time saved increases as it goes up, which indicates an increased benefit from a larger number of impacted buses. By considering more buses in the model, the algorithm will give a wider systematic decision. On the other hand, the marginal benefits of time saving decreases as the number of buses increases -- on average it is less than 1% when the number of buses is greater than four. And having more buses will inevitably increase the computational effort, which is less desirable in real-time application. This finding is generally consistent with the result from Eberlein et al. (2001), in which the authors found that the ideal number of impacted vehicles is three, and the change in benefits is relatively small beyond three. Having more buses also makes prediction of downstream travel time less realistic. Shen and Wilson (2001) pointed out that in practice the impact sets should not be too large because in the stochastic environment disruption drop quickly with further down stops and therefore including too much downstream stops in the impact set will give an overestimated result. Furthermore, we must consider data availability when including more impacted buses and stops in the implementation. If ***m*** and ***n*** are chosen too large, many input data are based on prediction, not on actual observations, which will cause inaccuracies in the holding strategy. Lastly, it is practically unnecessary to consider too many buses at once for a real-time single control point implementation. Therefore, in practice m={2, 3, 4} and n={3, 4} is appropriate for implementation.

[Table 7]
[Table 8]

In the second sensitivity study, we compared the performance of our holding strategy with others in the literature under different dispatching headways and demand levels. A similar experiment was described in detail in Fu and Yang (2002) and Fu et al. (2003) for testing the performance of a stop-skipping strategy for bus bunching problems in the Waterloo area. Although it is not about a bus holding strategy, we decided to compare our model with theirs because that is the only study we could find in the literature that gave a detail description about the simulation data set, based on which we were able to re-create the data and scenarios for comparison purposes. The running time has a large variation with coefficients of variation ranging from 0.21 to 0.53 than the CTA data (with the value range of 0.12-0.31) does, which means more bunching might occur on that route. To compare the performance of the strategies, we borrowed the parameter settings from Fu's paper, including the running time profile, the base demand profile, and the calibrated dwell time function. And we then use simulation to generate dispatching time with different dispatching headways.

To test the sensitivity of bus holding performance to dispatching headway, five scenarios were simulated with five different bus dispatching headways {3, 4, 5, 10, 15} minutes. The passenger demand was generated following the same procedure as described in Fu et al. (2003). In Figure 4, we compare the time savings for the stop skipping strategy from Fu et al. (2003) and the holding strategy based on our model. The benefit of stop-skipping decreases as the dispatching headway increases while the benefit of our holding strategy increases. This pattern suggests that the holding control is more appropriate for routes with a relatively longer headway than for those with a shorter headway.

To test the sensitivity to passenger demand, four scenarios were simulated with four different demand levels: base demand, 1.5x (base demand), 1.8 x (base demand), and 2x (base demand). Figure 5 shows that time saving increases smoothly and monotonically as the demand increases for our holding strategy, while for stop skipping strategy, there is an optimal level of demand at which the time saving is maximized, and beyond that time saving starts to decrease. This suggests that our holding strategy achieves more stable headways than the stop-skipping one even under different levels of passenger demand.

[Figure 4]
[Figure 5]

## 5 Conclusions

This study has presented a holding strategy of groups of buses at multiple control points. First we formulated the problem of holding a group of buses at a predefined control point. Considering possible passenger boarding during holding, the model was formulated as a non-convex optimization program with linear constraints and a heuristic is developed to help solve the problem. Furthermore, the algorithm was implemented at multiple control points with stochastic running times taken into consideration. The model was evaluated with a simulation case study by using the real-time bus operation data from CTA.

By simulating the multiple holding strategy based on the CTA real-time operation data, our proposed holding strategy has demonstrated benefits in reducing waiting time and improving the bus headway performance. The fast computing time of the proposed algorithm makes it possible for online implementation.

Comparing the proposed model with the traditional model which does not consider boarding activity while holding, we find that our model improves the downstream headway stability and outperforms the traditional model in that regard, especially when the control point is set at the stop with higher demand and when more impacted stops are included.

Sensitivity analysis shows that more impacted stops and buses in the model would improve performance. In practice, however, our study, consistent with others in the literature, seems to suggest up to four impacted buses and stops be sufficient to be considered without increasing unnecessary computational effort and decreasing prediction accuracy. Furthermore, based on our comparison analysis, bus holding is more appropriate than stop-skipping for larger dispatching headway and high demand routes.

There are several limitations of our model that are worth noting. These limitations also offer directions of further investigation in bus holding research. Firstly, the deterministic running time assumption in calculating holding times is strong: in reality, running time is uncertain. Although we use stochastic running times in the implementation, relaxing the assumption in the model would enable more accurate calculation of optimum holding times. Secondly, in implementation, the selection of the control point(s) is critical and must be investigated further in future work. Major transfer stops must be taken into consideration in the control point selection.

For example, if a major transfer point is downstream of a control point, large amount of delays may occur due to holding. Thirdly, a holding strategy is intended to reduce bus bunching by inserting slack into the bus schedule. On the other hand, when a bus is falling behind schedule, holding does not help with schedule adherence. Therefore, in practice holding strategies may be combined with other control strategies, e.g., stop skipping and adjusting bus cruising speed, in order to achieve a better performance within an integrated model.

## References


Abkowitz, M., Engelstein, I. (1984). Methods for maintaining transit service regularity. *Transportation Research Record.* Vol. 961, pp.1-8.

Abkowitz, M., Lepofsky, M. (1990) Implementing Headway-Based Reliability Control on Transit Routes. *ASCE, Journal of Transportation Engineering*, Vol. 1, No. 7, pp. 49-63.

Abkowitz, M., Tozzi,T. (1986) Transit Route Characteristics and Headway-Based Reliability Control. *Transportation Research Record*, No. 1078, pp. 11-16.

Barnett, A.(1974) On Controlling Randomness in Transit Operations. *Transportation Science* 8(2) ,pp. 102-116.

Bellei, G., Gkoumas, K. (2009). Threshold- and information-based holding at multiple stops, *IET Intelligent Transport Systems Journal*, Vol. 3, No. 3, pp. 304-313.

Bellei, G., Gkoumas, K. (2010) Transit vehicles' headway distribution and service irregularity. *Public Transport* Vol. 2, No. 4, pp. 269–289.

Cortes, C.E., Saez, D., Milla, F., Munez, A. and Riquelme, M. (2010). Hybrid Predictive Control for Real-time Optimization of Public Transport System's Operations Based on Evolutionary Multi-objective Optimization. *Transportation Research, Part C.* Vol. 18. No.5, pp757-769.

CTAPerformance Metrics 2011, available at http://www.transitchicago.com/assets/1/performance_metrics/CTA_Company_Wide_KPI_Dec_11.pdf, last accessed on Jan 5, 2012.

Delgado, F., Muñoz, J.C., Giesen, R., Cipriano, A. (2009) Real-Time Control of Buses in a Transit Corridor Based on Vehicle Holding and Boarding Limits. *Transportation Research Record.* Vol. 2090, pp. 59-67.

Desaulniers,G., Hickman, M. (2007). Public Transit. In *G.Laporte&C. Barnhart (Eds.), Handbooks in Operations Research and Management Science. Vol: 14, pp. 69-127.*

Ding, Y., Chien, S. (2001) Improving Transit Service Quality and Headway Regularity with Real Time Control. *Journal of the Transportation Research Board* , 1760. pp. 160-171.

Eberlein, X.J., Wilson, N.H.M., Bernstein, D. (2001) The Holding Problem with Real-time Information Available. *Transportation Science* 35(1), pp. 1-18.

Fu, L., Yang, X. (2002)Design and Implementation of Bus-holding Control Strategies with Real-Time Information. *Transportation Research Record.* Vol. 1791, pp. 6-12.

Fu, L., Liu, Q. and Kalamai, P. (2003) Real-Time Optimization Model for Dynamic Scheduling of Transit Operations. *Transportation Research Record.* Vol. 1857, pp. 48-55.

Herman, R.,Lam, T. (1974) Trip Characteristics of Journeys to and From Work, *Transportation and Traffic Theory, Proceedings of the Sixth International Symposium on Transportation and Traffic Theory, Elsevier: New York.*

Hickman, M. (2001) An Analytical Stochastic Model for the Transit Vehicle Holding Problem. *Transportation Science*35(3), pp. 215-237.

Lin, J., Ruan, M. (2009) Probability-based Bus Headway Regularity Measure. *ITE Intelligent Transport Systems*, Vol. 3, Iss. 4, pp. 400-408.

O'Dell, S., Wilson, N.H.M (1999). Optimal Real-Time Control Strategies for Rail Transit Operations during Disruption. *Computer Aided Systems in Public Transport.* Lecture Notes in Economics and Mathematical Systems Vol.471, pp. 299–323.

Oort, N.V, Wilson,N.H.M., Nes, R.V.(2010). Reliability Improvement in Short Headway Transit Services. *Transporation Research Record. Volume: 2143, pp. 67-76.*
Puong, A., Wilson, N.H.M.(2008). A Train Holding Model for Urban Rail Transit Systems. *Computer Aided Systems in Public Transport.* Lecture Notes in Economics and Mathematical Systems Vol. 600, pp. 319-337.

Shen, S., Wilson, N.H.M. (2001). An Optimal Integrated Real-time Disruption Control Model for Rail Transit Systems.*Computer-Aided Scheduling of Public Transport*, Lecture Notes in Economics and Mathematical Systems *Vol. 505,* pp. 335-363.

Sun, A., Hickman, M. (2008) The Holding Problem at Multiple Holding Stations. *Computer-aided Systems in Public Transport,,* Volume 600, Part III, pp. 339-359.

TCRP Report 100: Transit Capacity and Quality of Service Manual, 2nd Edition. 2003.

Toledo, T., Cats, O., Burghout, W. and Koutsopoulos, H.N. (2010) Mesoscopic simulation for transit operations*Transportation Research Part C* .Volume: 18, Issue: 6, pp. 896-908.

Turner, J.K. , J.G. Wardrop (1951) The Variation of Journey Time in Central London, *Road Research Laboratory Note No. RN/1511/JKT.JGW.*

Turnquist, M.A. (1981) Strategies for Improving Reliability of Bus Transit Service. *Transportation Research Record,* No. 818, pp. 7-13.

Welding, P.I. (1957) The Instability of a Close-interval Service. *Operational Research Quarterly* 8 (3), pp.133-142.

Xuan, Y., Argote, J., Daganzo, C.F.(2011). Dynamic Bus Holding Strategies for Schedule Reliability: Optimal Linear Control and Performance Analysis. *Transportation Research Part B. Vol* 45.pp1831-1845.

**Table 1** Convergence analysis for selected sample runs

| **ε=1.0e-5** | **Absolute differences, sample run 1-6** | | | | | |
|---|---|---|---|---|---|---|
| **Problem Size (m, n)** | **(4,3)** | | **(10,10)** | | **(6,8)** | |
| **X (no. iteration)** | **Y1** | **Y2** | **Y3** | **Y4** | **Y5** | **Y6** |
| 1 | 1.94E+01 | 2.31E-01 | 1.15E+02 | 1.66E+01 | 9.43E-01 | 2.64E+01 |
| 2 | 2.90E-01 | 8.83E-04 | 7.73E-01 | 8.08E-02 | 8.26E-03 | 5.17E-01 |
| 3 | 4.33E-03 | **3.53E-06** | 4.48E-03 | 7.12E-04 | 7.24E-05 | 1.95E-03 |
| 4 | 6.48E-05 | 1.69E-08 | 2.71E-05 | **4.84E-06** | **6.41E-07** | 7.26E-05 |
| 5 | **9.69E-07** | 1.12E-10 | **2.64E-07** | 4.32E-08 | 6.22E-09 | **3.06E-06** |
| 6 | 1.45E-08 | 0.00E+00 | 3.19E-09 | 2.86E-10 | 0.00E+00 | 1.43E-07 |
| 7 | 2.16E-10 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 6.67E-09 |
| 8 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 0.00E+00 | 3.34E-10 |
| Iteration required | 5 | 3 | 5 | 4 | 4 | 5 |
| CPU time (sec) | 0.1959 | 0.1842 | 0.8143 | 0.7108 | 0.514 | 0.6322 |

**Table 2.** Running time profile for the selected CTA route

| ***j*** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $mR_j$ | 5.151 | 5.130 | 12.322 | 3.415 | 3.389 | 9.549 | 9.218 | 9.260 | 6.429 | 2.982 | 2.336 |
| $varR_j$ | 1.293 | 1.340 | 1.755 | 0.196 | 0.685 | 1.957 | 2.062 | 2.269 | 1.037 | 0.896 | 1.046 |

**Table 3.** Summary of the single control point problem

| Run # | Impacted Bus group | **Inputs** | | | **Results** | | | |
|---|---|---|---|---|---|---|---|---|
| | | Arrival time | Departure time | Bus Loading | Holding time (min) | Time saving (%) | CPU time | #iterations |
| 1 | {2,3,4} | $a_{2,7}$, $a_{3,7}$, $a_{4,7}$ | $d_{1,7}$, $d_{1,8}$, $d_{1,9}$, $d_{1,10}$ | $L_{2,6}$, $L_{3,6}$, $L_{4,6}$ | 5.961<br>6.873<br>8.728 | 7.13 | 0.374 | 6 |
| 2 | {3,4,5} | $a_{3,7}$, $a_{4,7}$, $a_{5,7}$ | $d_{2,7}$ $d_{2,8}$, $d_{2,9}$, $d_{2,10}$ | $L_{3,6}$, $L_{4,6}$, $L_{5,6}$ | 3.309<br>7.990<br>4.983 | 8.97 | 0.125 | 7 |

**Table 4.** Comparison of multiple holding and single holding strategies

| Strategy | Control point | Average holding time (min) | Time saving (%) | Improved headway stability at downstream 5 stops (%) | | | | |
|---|---|---|---|---|---|---|---|---|
| Multiple | 3 | 3.82 | 7.3 | 32.82 | 31.09 | 13.69 | 10.01 | 6.3 |
| holding | 7 | 5.01 | 13.65 | 65.44 | 47.34 | 33.01 | 14.31 | 9.16 |
| Single holding | 7 | 5.35 | 13.29 | 55.42 | 36.26 | 23.88 | 13.63 | 13.06 |

**Table 5.** Location of Control Points

| Stops<br>Control points | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| arrival rate (person/min) | 2.94 | 1.04 | 2.52 | 2.09 | 1.41 | 1.39 | 3.07 | 1.33 | 1.93 | 1.52 | 2.04 | 1.92 | 0 |
| n=2 | | Y | | | Y | | | Y | | | Y | | |
| n=3 | | | Y | | | | Y | | | | Y | | |
| n=4 | | Y | | | | | Y | | | | | | |
| n=5 | | Y | | | | | | Y | | | | | |

**Table 6.** Comparison between Proposed Model and Traditional Model

| | | | Proposed | Traditional | difference between P and T | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Scenario | control points | arrival rates (passenger/min) | average holding time(min) | average holding time(min) | Difference in total waiting time saving (passenger.min) | Difference in improved headway stability(%) | | | | |
| | 2 | 1.04 | 1.48 | 1.68 | 6.33 | 2.48 | 1.13 | | | |
| | 5 | 1.41 | 1.52 | 1.76 | 10.64 | 4.69 | 3.17 | | | |
| | 8 | 1.33 | 4.97 | 5.33 | 17.69 | 4.86 | 2.64 | | | |
| n=2 | 11 | 2.04 | 2.01 | 2.11 | 15.36 | 3.14 | 2.22 | | | |
| | 3 | 2.52 | 3.82 | 4.01 | 11.45 | 6.86 | 5.57 | 3.68 | | |
| | 7 | 3.07 | 5.01 | 5.82 | 27.42 | 8.07 | 6.41 | 3.28 | | |
| n=3 | 11 | 2.04 | 2.06 | 2.13 | 19.94 | 3.31 | 2.37 | 2.4 | | |
| | 2 | 1.04 | 1.35 | 1.44 | 9.25 | 4.97 | 8.55 | 1.58 | 0.44 | |
| n=4 | 7 | 3.07 | 6.16 | 7.21 | 43.31 | 10.03 | 3.62 | 3.05 | 1.14 | |
| | 2 | 1.04 | 1.31 | 1.43 | 11.35 | 2.45 | 1.63 | 0.45 | 0.29 | 0.12 |
| n=5 | 8 | 1.33 | 4.47 | 5.02 | 13.26 | 4.18 | 3.14 | 2.5 | 1.89 | 0.94 |

**Table 7.** Average holding time (min) by number of impacted buses and stops

| No. of impacted buses | No. of impacted stops | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 5.442 | 4.983 | 4.332 | 3.367 | 3.308 |
| 2 | 6.286 | 5.644 | 5.207 | 4.948 | 4.908 |
| 3 | 5.997 | 5.453 | 4.967 | 4.542 | 4.260 |
| 4 | 4.228 | 4.104 | 3.881 | 3.394 | 3.169 |
| 5 | 4.630 | 4.210 | 3.997 | 3.105 | 2.944 |

**Table 8.** Time Saving (%) by number of impacted buses and stops

| No. of impacted buses | No. of impacted stops | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | 9.98 | 12.42 | 13.08 | 13.86 | 13.92 |
| 2 | 11.75 | 12.08 | 13.71 | 13.84 | 13.91 |
| 3 | 12.61 | 13.52 | 14.09 | 14.64 | 14.65 |
| 4 | 12.92 | 13.38 | 14.49 | 15.81 | 15.95 |
| 5 | 13.06 | 13.93 | 14.84 | 15.72 | 15.63 |

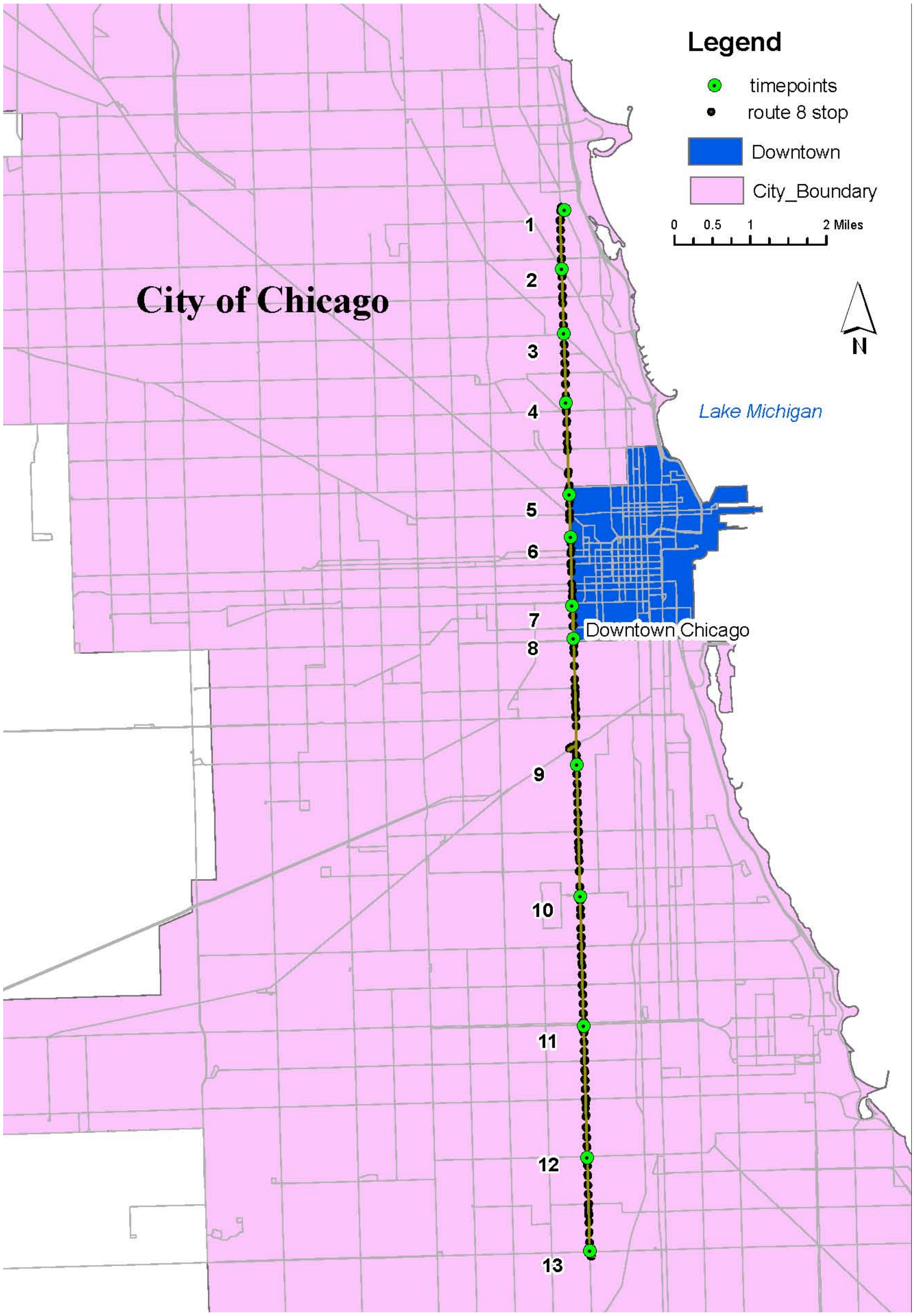



**Figure 1.CTA Bus system and the study route**

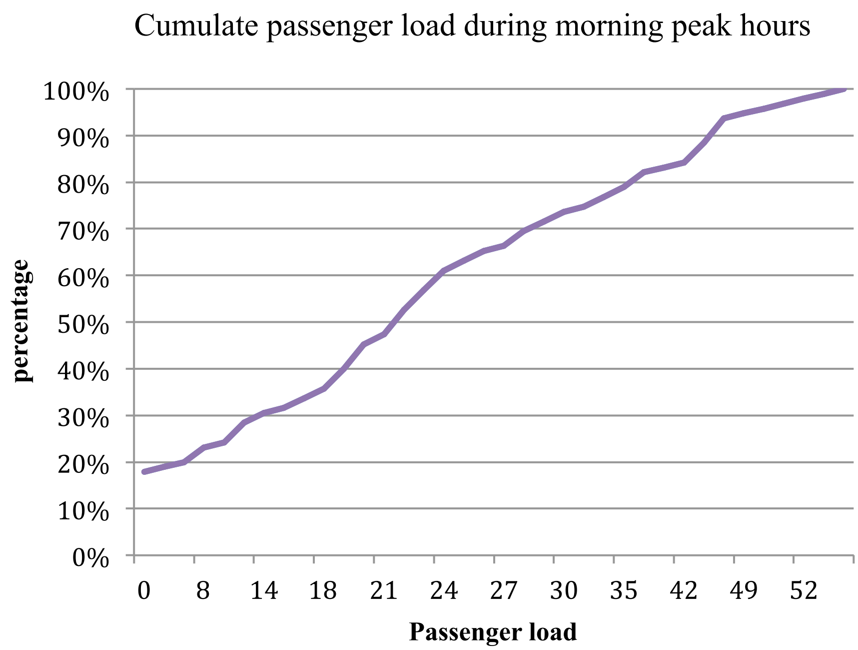
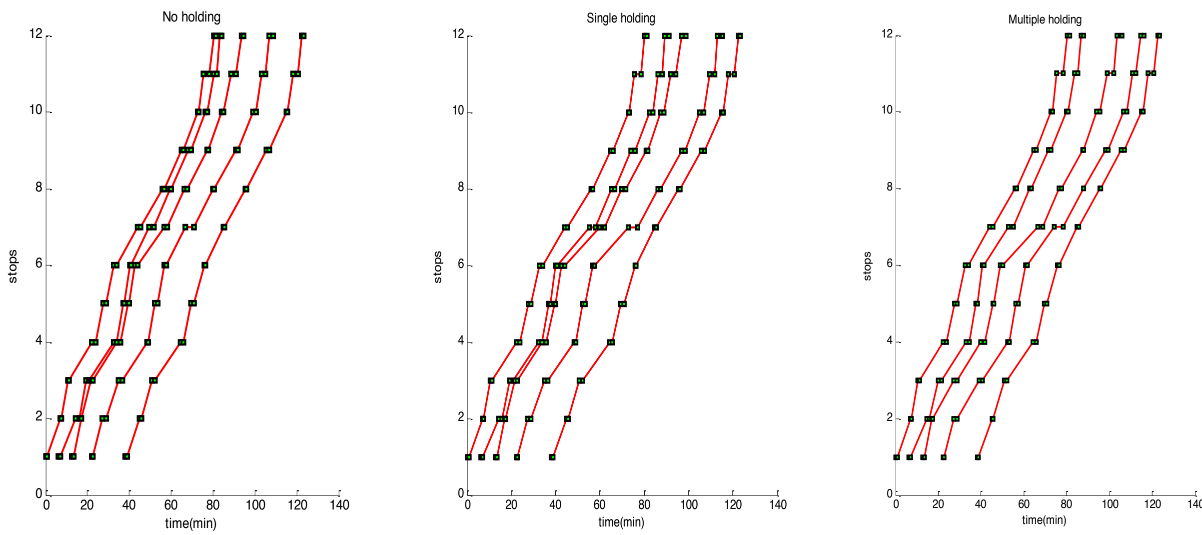



**Figure 2. Morning peak hour passenger load profile**




a. No holding b. Single holding c. Multiple holding

**Figure 3. Vehicle trajectories with/without holding strategy**

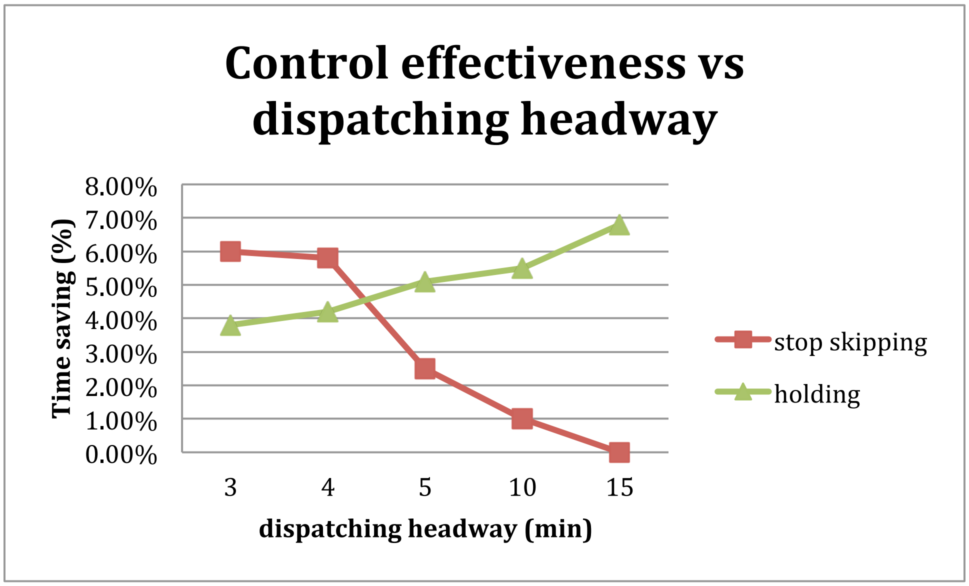



**Figure 4. Control effectiveness by dispatching headway**

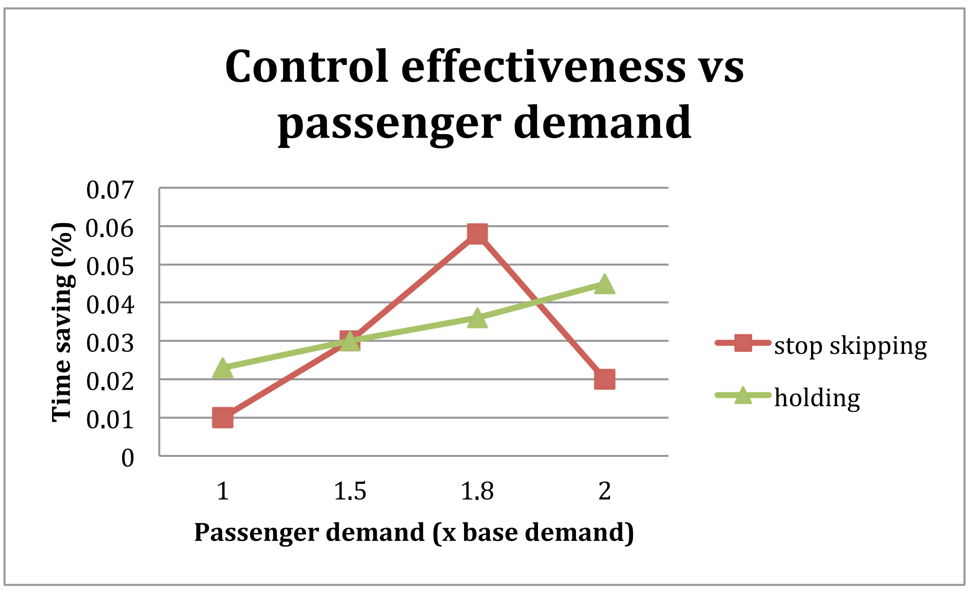



**Figure 5. Control effectiveness by passenger demand**